Les propriétés des nombres premiers, p. 99 de *Pour l'honneur de l'esprit humain* de Jean Dieudonné "*Les problèmes prolifiques*" (§ B de la section 2 du Chapitre IV. Quelques problèmes de mathématiques classiques).

B. Les propriétés des nombres premiers

À ma connaissance, dans aucune civilisation antique autre que la civilisation grecque, on n'avait songé avant le Ve siècle avant J.-C. à la décomposition d'un entier en facteurs premiers. Cette décomposition, que nous écrivons maintenant

$$n = p_1^{k_1} p_2^{k_2} \dots p_r^{k_r} \tag{11}$$

où $p_1, p_2, \dots, p_r$, sont des nombres premiers et les k_i des exposants au moins égaux à 1, n'apparaît pas explicitement chez Euclide faute de notations adéquates. Mais il démontre les trois propriétés suivantes (exprimées en langage moderne) :

a) Tout entier est premier ou divisible par un nombre premier (Livre VII, 31).

b) Si p est un nombre premier, une puissance p^m ne peut être divisible que par les nombres p^r avec $r < m$ (Livre IX, 13).

c) Si un nombre premier divise un produit ab de deux entiers et ne divise pas a, il divise b (Livre VII, 32).

À partir de là, il est facile, en raisonnant par récurrence, d'établir l'existence et l'unicité de la décomposition (11).

Rappelons que nous avons cité au chapitre II, § 5, le plus beau théorème de l'arithmétique grecque, le fait qu'il y a une infinité de nombres premiers. La démonstration donnée par Euclide est très simple (voir Hardy, [7], p. 28, et Appendice I) ; mais je préfère en présenter une autre, due à Euler, parce qu'elle a ouvert la voie à ce qu'on appelle la "théorie analytique des nombres premiers" ; on peut toutefois la présenter sans utiliser autre chose que l'algèbre élémentaire.

C'est un raisonnement "par l'absurde", où on suppose que

$$p_1 = 2, p_2 = 3, p_3 = 5, \dots, p_r$$

(rangés par ordre croissant) sont les seuls nombres res premiers, et on va en déduire une conclusion absurde. Par (11), tout entier n se décomposerait en le produit du second membre d'une seule manière, en admettant cette fois que certains des exposants k_i peuvent être 0 (le facteur $p_i^{k_i}$ étant alors alors remplacé par 1). Prenons un entier N arbitrairement grand et considérons le produit

$$S_{N,r} = \left(1 + \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^N}\right) \times \left(1 + \frac{1}{3} + \frac{1}{3^2} + \dots + \frac{1}{3^N}\right) \dots \left(1 + \frac{1}{p_r} + \frac{1}{p_r^2} + \dots + \frac{1}{p_r^N}\right). \tag{12}$$

Pour effectuer ce produit, il faut prendre un terme dans chaque parenthèse et en faire le produit, puis faire la somme de tous ces partiels s'écrivent $1/n$ où n est de la forme (11), mais avec la restriction que $0 \leq k_i \leq N$ pour *tous* les exposants. Par l'unicité de la décompositiontous ces produits

Transcription en LaTeX : Denise Vella-Chemla, février 2026.

partiels sont différents ; mais la remarque essentielle est que tous les entiers n compris entre 1 et 2^N apparaissent (une seule fois d'après ce qui vient d'être dit) dans un produit partiel $1/n$. En effet, si $1 \leq n \leq 2^N$, dans la décomposition (11) aucun des exposants $k_2 k_3, \ldots, k_r$ ne peut être plus grand que $N-1$, sans quoi le nombre n serait au moins égal à 3^N alors qu'on l'a supposé $\leq 2^N$; $1/n$ apparaît donc bien comme un des produits partiels dans l'expression de $S_{N,r}$. Bien entendu, il y a dans $S_{N,r}$ d'autres produits partiels, mais ce que l'on a montré c'est que l'on a

$$S_{N,r} \geq 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \ldots + \frac{1}{2^N - 1} + \frac{1}{2^N}. \tag{13}$$

La somme du second membre n'est pas facile à évaluer, mais on peut la remplacer par un nombre plus petit de la façon suivante : on groupe les termes en sommes partielles s'arrêtant aux puissances de $\frac{1}{2}$:

$$1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^{N-1}+1)} + \frac{1}{2^{N-1}+2} + \ldots + \frac{1}{2^N}\right)$$

La première parenthèse a 2 termes au moins égaux à $1/4$, donc elle est $\geq \frac{1}{2}$. La seconde a 4 termes au moins égaux car elle a 2^{k-1} termes au moins égaux à $1/8$ donc elle est $\geq 4/8 = \frac{1}{2}$. Continuant ainsi, on voit que chaque parenthèse est $\geq \frac{1}{2}$, car elle a 2^{k-1} termes et tous sont $\geq 1/2^k$. Comme il y a $N-1$ parenthèses, on a finalement

$$S_{N,r} \geq 1 + \frac{1}{2}N. \tag{14}$$

Mais on peut exprimer $S_{N,r}$ autrement par la formule donnant la somme d'une progression géométrique

$$1 + a + a^2 + \ldots + a^N = (1 - a^{N+1})/(1-a)$$

d'où

$$S_{N,r} = \frac{\left(1 - \frac{1}{2^{N+1}}\right)\left(1 - \frac{1}{3^{N+1}}\right)\ldots\left(1 - \frac{1}{p_r^{N+1}}\right)}{\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\ldots\left(1 - \frac{1}{p_r}\right)} \tag{15}$$

Si on remplace tous les facteurs du numérateur par 1, on a une inégalité

$$S_{N,r} \leq \frac{1}{\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\ldots\left(1 - \frac{1}{p_r}\right)} = A_r \tag{16}$$

et le second membre ne dépend plus de N. Comparant (14) et (16) on obtient $1 + \frac{1}{2}N \leq A_r$ où $N \leq 2(A_r - 1)$ et comme N est aussi grand qu'on on est arrivé à une absurdité.

L'idée d'Euler, qui a été le germe de tous les progrès ultérieurs, fut de remplacer dans (12) les inverses $1/p_j$ des nombres premiers par une puissance $(1/p_j)^m$ où l'exposant est plus grand que 1 (mais non nécessairement entier). Mais la formule que l'on obtient ainsi ne peut malheureusement

se décrire qu'en utilisant des notions d'analyse, séries et produits infinis, que nous ne pouvons employer ici (voir Appendice II).

Une fois acquis le fait que la suite des nombres premiers ne s'arrête pas, on peut du moins faire des tables donnant les nombres premiers inférieurs à un certain nombre. On a eu assez tôt des tables qui allaient jusqu'à $3 \cdot 10^6$ (trois millions), et les ordinateurs peuvent faire beaucoup mieux. La plus ancienne méthode connue pour fabriquer ces tables est ce qu'on appelle le "crible d'Ératosthène". Pour avoir les nombres premiers $\leq x$, on écrit la suite de tous les entiers $2, 3, 4, 5, \ldots, x$; on barre les multiples de 2 à partir de 4, puis les multiples de 3 à partir de 6, les multiples de 5 à partir de 10, et ainsi de suite : de façon précise, après la k-ième opération, les $k+1$ plus petits nombres non barrés sont premiers, et si p_{k+1} est le plus grand d'entre eux, la $(k+1)$-ième opération consiste à barrer les multiples de $p_{k+1<}$ à partir de $2p_{k+1}$. On peut s'arrêter au dernier nombre premier p_r qui est $\leq \sqrt{x}$; en effet, si un entier m est tel que $\sqrt{x} < m \leq x$ et est non barré, il ne peut être un produit ab avec $a > 1, b > 1$, car un au moins des nombres a, b serait $\leq \sqrt{x}$, donc divisible par un des nombres premiers déjà trouvés, et m aurait dû être barré. Ainsi, tous les nombres non barrés et $> \sqrt{x}$ sont premiers.

Il y a des méthodes plus puissantes pour établir des tables de nombres premiers ; mais le crible d'Eratosthène est peut-être le procédé qui a suggéré à Euler de considérer le nombre A_r de la formule (16) : en effet, pour un nombre premier $p \leq \sqrt{x}$, il y a $\left(1 - \frac{1}{p}\right) x$ nombres tels que $1 \leq m \leq x$ et qui ne sont pas multiples de p. S'il n'y avait pas dans le "crible d'Ératosthène" des nombres qui sont barrés plusieurs fois, il y aurait à peu près

$$\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right) \ldots \left(1 - \frac{1}{p_r}\right) x \tag{17}$$

nombres premiers inférieurs à x, où p_r est le plus grand nombre $\leq \sqrt{x}$; le facteur de x dans (17) est précisément le dénominateur du nombre A_r. Il dépend naturellement de x, et Euler put montrer qu'il tend vers 0 lorsque x croît indéfiniment. Cependant, l'examen d'une table de nombres premiers fait apparaître une extrême irrégularité dans leur distribution ; on connaît un grand nombre de nombres premiers p tels que $p + 2$ soit aussi premier [1] (on dit qu'ils forment une paire de nombres premiers "jumeaux") et on soupçonne même qu'il y en a une infinité, bien qu'actuellement, on ne sache pas le démontrer. D'autre part, il y a dans la suite des entiers des "trous" aussi grands qu'on veut où il n'y a aucun nombre premier, par exemple la suite

$$n! + 2, \qquad n! + 3, \qquad \ldots, \qquad n! + n.$$

Même Euler était découragé par ces tables et pensait que la répartition des nombres premiers était "un mystère auquel l'esprit humain ne saurait jamais pénétrer". Mais, à la fin du XVIII$^{\text{e}}$ siècle, Legendre et Gauss, indépendamment, eurent l'idée qu'en moyenne, la répartition des nombres premiers obéissait à des lois simples. Gauss considérait tous les nombres premiers entre un nombre x et le nombre $x + 1000$; si $N(x)$ est ce nombre, il observa sur les tables que lorsque x est grand, le rapport $N(x)/1000$ est voisin de $1/\log x$. S'il s'agissait d'une couche continue de matière répartie

1. En 1985, on connaissait 3 424 506 nombres $p \leq 10^9$ tels que p et $p + 2$ soient premiers.

sur une droite et de densité $1/\log x$, la quantité comprise entre 2 et x serait

$$\mathrm{li}(x) = \int_2^x \frac{\mathrm{d}t}{\log \mathrm{t}}$$

qu'on appelle le logarithme Gauss pensa donc que le nombre $\pi(x)$ des nombres premiers compris entre 2 et x devait être "approché" par le nombre $\mathrm{li}(x)$.

On précisa cette idée en conjecturant que lorsque x croît indéfiniment, le rapport $\pi(x)/\mathrm{li}(x)$ tend vers 1 ; c'est ce qu'on appela le "théorème des nombres premiers" ; si $p_1, p_2, \ldots, p_n$ est la suite croissante des nombres premiers, on montre que ce théorème équivaut à dire que le rapport $p_n/n\log n$ tend vers 1.

L'examen des tables montre que ces conjectures sont vraisemblables ; par exemple, pour $x = 4\cdot 10^{16}$, on a

$$\pi(x) = 1\ 075\ 292\ 778\ 753\ 150, \qquad \mathrm{li}(x) - \pi(x) = 5\ 538\ 861.$$

Pendant tout le XIX$^{\mathrm{e}}$ siècle, beaucoup de mathématiciens s'attelèrent à la démonstration du théorème des nombres premiers. Mais il ne fut prouvé qu'en 1896, presque simultanément par J. Hadamard et C. de la Vallée-Poussin. Malheureusement, toutes les méthodes de démonstration reposent sur l'étude d'une fonction notée $\zeta(s)$ introduite par Riemann, et qui exige des techniques avancées d'Analyse, que nous ne pouvons songer à décrire.

Mais les mathématiciens ne se sont pas contentés de ce succès ; ils voudraient savoir comment se comporte la différence

$$\pi(x) - \mathrm{li}(x)$$

Une conjecture de Riemann sur les propriétés de sa fonction $\zeta(s)$ entraînerait, si elle était vraie, que

$$\frac{\pi(x) - \mathrm{li}(x)}{x^{\frac{1}{2}+\alpha}}$$

tend vers 0 pour tout exposant $\alpha > 0$. Malheureusement, malgré 130 ans d'efforts, personne n'a encore pu prouver ni infirmer l'hypothèse de Riemann, qui reste un des problèmes ouverts les plus importants des mathématiques, parce que sa résolution entraînerait de grands progrès dans de nombreuses parties de la théorie des nombres.

On a longtemps cru que l'on a toujours $\pi(x) < \mathrm{li}(x)$; les tables montrent que c'est vrai pour $x \leq 10^8$. Mais Littlewood a établi qu'il y a une infinité d'entiers x tels que

$$\pi(x) - \mathrm{li}(x) > \frac{\sqrt{x}}{2\log x}$$

et aussi une infinité d'entiers x tels que

$$\pi(x) - \mathrm{li}(x) < -\frac{\sqrt{x}}{2\log x}$$

On ne connaît pas encore la valeur du plus petit entier x_0 pour lequel $\pi(x) - \mathrm{li}(x)$ change de signe, mais il est certainement très grand. Ce résultat confirme évidemment l'impression d'extrême irrégularité dans la distribution locale des nombres premiers.

En 1785, Legendre, en vue d'applications à la théorie des formes quadratiques, eut besoin d'une précision au théorème d'Euclide, à savoir que si a et b sont deux entiers premiers entre eux, il y a une infinité de nombres premiers dans la progression arithmétique $an + b$. Dans certains cas particuliers, cela se voit aisément en généralisant convenablement la méthode d'Euclide, par exemple pour les progressions arithmétiques $4n + 3$ et $6n + 5$ (Appendice 1). Mais la preuve du théorème général n'a jamais être obtenue par des moyens élémentaires ; elle fut donnée par Dirichlet en 1837, par l'utilisation de fonctions qui généralisent la fonction $\zeta(s)$ de Riemann.

On note $\pi(x; a, b)$ le nombre de nombre premiers au plus égaux à x dans la progression arithmétique $an + b$; Hadamard et de la Vallée Poussin ont étendu leurs méthodes pour obtenir une estimation de $\pi(x; a, b)$; ils ont montré que le rapport

$$\frac{\pi(x; a, b)}{\dfrac{x}{\log x}}$$

a pour limite $1/\varphi(a)$, ù $\varphi(a)$ est le nombre des nombres m premiers à a et tels que $0 < m < a$; cette formule avait déjà été conjecturée par Legendre.

Appendice I

Les nombres premiers de la forme $4k - 1$ *ou* $6k - 1$

Un nombre premier autre que 2 est nécessairement de l'une des formes $4k+1$ ou $4k-1$ avec $k > 1$; le raisonnement d'Euclide, légèrement modifié, montre qu'il y a une infinité de nombres premiers de la forme $4k - 1$.

On considère la suite croissante de nombres premiers de la forme $4k - 1$

(1) $$p_1 = 3 < p_2 = 7 < \ldots < p_r$$

On forme le nombre

(2) $$N = 4p_1p_2 \ldots p_r - 1,$$

Il ne peut évidemment être divisible par aucun des nombres de la suite (1) ; d'autre part, il ne peut être produit de puissances de nombres premiers de la forme 4k + 1 , car tout produit

$$(4a + 1)(4b + 1) = 4(4ab + a + b) + 1$$

est encore de la forme $4c+1$, donc aucun produit de facteurs premiers de la forme $4k+1$ ne peut être égal au nombre N défini par (2). Un des facteurs premiers de N est donc de la forme $p = 4k - 1$, et pest distinct des nombres de la suite (1). Tout nombre premier autre que 2 et 3 est nécessairement

de l'une des formes $6k+1$ ou $6k-1$; en remplaçant 4 par 6 dans le raisonnement précédent, on montre qu'il y a une infinité de nombres premiers de la forme $6k-1$.

Appendice II.

La décomposition de $\zeta(s)$ en produit eulérien

Soit $p_1 < p_2 < \ldots < p_r < \ldots$ la suite croissante (infinie) des nombres premiers. Soit d'autre part s un exposant plus grand que 1. La formule d'Euler sur les nombres premiers est la suivante :

$$(1) \qquad \frac{1}{\left(1-\frac{1}{p_1^s}\right)\left(1-\frac{1}{p_2^s}\right)\ldots\left(1-\frac{1}{p_r^s}\right)\ldots} = 1+\frac{1}{2^s}+\frac{1}{3^s}+\ldots+\frac{1}{n^s}+\ldots$$

où au second membre figurent tous les entiers, chacun une fois. Pour donner un sens à cette formule, il faut prouver trois choses :

I) Si, pour tout entier m, on pose

$$F(m) = 1+\frac{1}{2^s}+\frac{1}{3^s}+\ldots+\frac{1}{m^s}$$

la suite $(F(m))$ a une limite S.

II) Si on pose

$$G(r) = \frac{1}{\left(1-\frac{1}{p_1^s}\right)\left(1-\frac{1}{p_2^s}\right)\ldots\left(1-\frac{1}{p_r^s}\right)}$$

la suite $(G(r))$ a une limite P.

III) On a $P = S$.

Preuve de I : Si $m < m'$, on a évidemment $F(m) < F(m')$. Comme au § 2, B), on considère différence $F(2^N) - F(2^M)$ pour $M < N$, qui s'écrit

$$\left(\frac{1}{(2^M+1)^s}+\ldots+(\frac{1}{2^{(M+1)s}}\right)+\ldots+\left(\frac{1}{(2^{N-1}+1)^s}+\ldots+(\frac{1}{2^{Ns}}\right).$$

Dans chaque parenthèse,

$$\frac{1}{(2^k+1)^s}+\frac{1}{(2^k+2)^s}+\ldots+\frac{1}{2^{(k+1)s}}$$

il y a 2^k termes, et chacun est plus petit que $\frac{1}{2^{ks}}$, donc la somme des termes de la parenthèse est plus petite que $\frac{1}{(2^{s-1})^k}$. Par suite $F(2^N) - F(2^M)$ est plus que la somme de la progression géométrique de premier terme $\frac{1}{(2^{s-1})^M}$ et de raison $\frac{1}{(2^{s-1})^s}$ soit

$$(2) \qquad F(2^N) - F(2^M) \leq \frac{a}{(2^{s-1})^M} - \frac{a}{(2^{s-1})^N}$$

où

$$a = \frac{1}{1-\frac{1}{2^{s-1}}}$$

On a ainsi formé des intervalles $\left[F(2^M), F(2^M) + \frac{a}{(2^{s-1})^M}\right]$ qui sont emboîtés, de longueur tendant vers 0 et qui contiennent tous les nombres $F(m)$ pour $m > 2^M$; le principe de Cauchy montre que la limite S de la suite $(F(m))$ existe, l'unique nombre commun à tous ces intervalles, et on a $F(m) < S$ pour tout entier m.

Preuve de II) et III) : Si $r < r'$, on a $G(r) < G(r')$. Pour chaque entier N, il y a un entier r_N tel que, pour tous les entiers $n \leq 2^N$, la décomposition de n en facteurs premiers ne contient que les nombres premiers $p_1, p_2, \ldots, p_{r_N}$; le raisonnement de § 2. B) montre que l'on a $F(2^N) \leq G(r_N)$. D'autre part, pour chaque entier r fixé, $G(r)$ est limite, pour M croissant indéfiniment, de

$$\frac{\left(1 - \frac{1}{p_1^{sM}}\right)\left(1 - \frac{1}{p_2^{sM}}\right)\ldots\left(1 - \frac{1}{p_r^{sM}}\right)}{\left(1 - \frac{1}{p_1^{s}}\right)\left(1 - \frac{1}{p_2^{s}}\right)\ldots\left(1 - \frac{1}{p_r^{s}}\right)}$$

et ce nombre est au plus $F((p_1 p_2 \ldots p_r)^M) < S$. Pour tout $r > r_N$, tous les nombres $G(r)$ sont donc dans l'intervalle

$$[F(2^N), S]$$

et le principe de Cauchy montre que la suite $(G(r))$ a une limite P égale à S.

Le nombre S, valeur commune des deux membres de (1), se note $\zeta(s)$.